Performance Analysis of Steiner Tree-based Decentralization Mechanism (STDM) for Privacy Protection in Wireless Sensor Networks

B. Sathish Babu[1], Jayashree N[2], and Pallapa Venkataram[3]
(Corresponding author: B. Sathish Babu)

Department of Computer Science ans Engineering[1]
Siddaganga Institute of Technology, Tumkur - 572103, Karnataka, India.
Department of Information Science and Engineering[2]
Ballari Institute of Technology and Management, Bellary - 583104, Karnataka, India.
Protocol Engineering and Technology Unit, Department of ECE[3]
Indian Institute of Science, Bangalore - 560012, Karnataka, India.
(Email: {bsb, pallapa}@ece.iisc.ernet.in, darecse@gmail.com)

(Received Nov. 26, 2011; revised and accepted Jan. 30, 2012)

Abstract

Privacy and Security of the data are the major concern in Wireless Sensor Networks (WSNs). Many applications which are based on WSN require data exchanges with data privacy intact of the sensed data. Using minimum tree structure can significantly reduce the number of nodes in data transmission. In this paper, we propose a privacy mechanism, STDM, based on Steiner tree and decentralization mechanism in order to provide privacy of the data with minimum number of hops to the sink for WSNs. Simulation results show that STDM performs efficiently compared to some of the existing approaches. It gives high path diversity which guarantees the increase in data privacy. This paper also analyses the performance of STDM with respect to metrics such as path diversity, energy consumption, reliability, communication overhead and computation costs.

Keywords: Communication overhead, computation cost, data privacy, decentralization, security, Steiner tree

1 Introduction

Data privacy is to share the data only with the trusted users and protecting the personally identifiable information. There are two types of attacks that may occur on the data being transmitted in WSN: (1) Active attacks where unauthorized attackers monitor, listen and modify the data stream in the communication channel; and (2) Passive attacks where unauthorized attackers monitor and listen to the data stream in the communication channel. Attacks against privacy come under passive attacks. WSN is vulnerable to security attacks due to the broadcast nature of the transmission medium and the nodes which are often placed in a hostile environment where they are not physically protected. The data privacy cannot be protected under such conditions. Hence, there is a need of a proper mechanism to be applied for WSN to protect the data against privacy attacks.

The attacks against privacy of data are [14, 24]: a) Monitoring and eavesdropping: an adversary can easily discover the communication contents by listening or snooping the data. The control information about the sensor network configuration conveyed by the traffic is more detailed compared with that of the location server and hence, results in an effective eavesdropping against privacy protection; b) Traffic analysis: an adversary can obtain information about the communication pattern just through the sensor activities in network resulting in the network breakdown; and c) Camouflages adversary: one can insert a malicious node or compromise the nodes to hide in the network and misroute the packets causing privacy problems.

Some of the mechanisms available for privacy protections are [6]: a) Anonymity mechanisms which depersonalizes the data before release using techniques like, decentralizing sensed data, changing data traffic and node mobility; b) Policy-based approach where access control decisions and authentications are made; and c) Information flooding [19] which is a single path routing including randomized data routing and phantom traffic generation mechanisms using techniques like: baseline flooding where the data is transmitted and received only once, probabilistic flooding where the subset of nodes participate in data transmission, phantom flooding which is to direct the data to different locations of the network avoiding adversary from getting a steady stream of data and flooding with fake messages.

1.1 Steiner Tree

A Steiner tree is defined as the minimum weight tree connecting a designated set of vertices called terminals, in an undirected, weighted graph or points in a space. The tree may include non-terminals called Steiner points or Steiner vertices and is a problem of least cost multicast routing to find the tree that spans a set of destinations with the minimum cost over all links [9]. Steiner tree connects the group members through a given graph by minimizing the used resources and is similar to Minimum Spanning Tree but differs as in Steiner tree intermediate vertices and edges may be added to the graph in order to minimize the length of the tree.

Steiner tree reduces the number of forwarders and constructs multiple trees in parallel with reduced number of common nodes among them, i.e., when there are multiple paths to be utilized from a source node to the destination node, multiple trees are constructed from source node through different neighbor nodes and with reduced number of common nodes among the tree constructed from these neighbor nodes. According to its definition, it reduces the number of nodes and links used for constructing a delivery tree. Hence, it is very useful in representing solution to multicast routing problems since it deals with minimizing the cost of multicast routing tree [21, 23].

1.2 Decentralization Mechanism

The decentralization privacy mechanism which divides and distributes the data along different paths from source to the sink is used for data privacy in STDM. The number of paths to the sink node from the source node depends upon the Steiner tree constructed and are distinguished with an identifier. All the nodes pre-share the same irreducible polynomial which is a large prime number. Since the original data is distributed among the nodes, an adversary when attacks, or compromises a node and obtains communication pattern, it is just of the path from that particular node and has no complete data. This controls privacy attacks to greater extent.

This process comes under secret sharing scheme [18], in which the original data is divided into n number of divisions and is distributed along the Steiner tree. The sending node sets a threshold number t without needing to consult the sink, i.e., original data can be recovered at the sink only if at least t divisions of data are received. This is called as threshold scheme (t,n) where $n \geq t$. The divisions of the data are encrypted and then forwarded from the source node along with a check code. In encryption phase, the data partitions are obtained using the Equation 1:

$$f(x) = D + a_1x + a_2x^2 + ... + a_{t-1}x^{t-1} \bmod \quad q \qquad (1)$$

where, D is the original data, $q \geq max(n, D)$ and $a_i(i = 1, 2, 3, ...n)$ are random images and q is sufficiently large prime number. Each share (u_i, v_i) can be obtained by substitution of the value $u_i(i = 1, 2, ...n)$ for $f(x)$, the values of q can be disclosed and the values of t,a_i and D are enclosed. To obtain the original data, there is a need of at least t partitions (u_i, v_i) to be collected at sink. When t divisions are collected, the original data can be decrypted using the Lagrange's interpolation method given in Equations 2 and 3.

$$D = \rho_1v_1 + \rho_2v_2 + ... + \rho_tv_t \qquad (2)$$

where,

$$\rho_j = \prod_{i=1, i\neq j}^{t} \frac{u_i}{u_i - u_j} mod \quad q. \qquad (3)$$

The accuracy of the obtained decrypted data is checked by using decrypted check code value. This approach achieves the following:

1) The decentralization of a data makes it easy to protect the original data from being accessed by an adversary. Even if an adversary attacks a node and captures data, it is just one of the partitions. Hence, a complete data cannot be obtained at any node which guarantees the data privacy. This reduces monitoring and Eavesdropping.

2) Data partitions are sent through different nodes to the destination. When an adversary tries to obtain the transmission path information, it gets only the data about that particular path which is incomplete. This reduces traffic analysis.

1.3 Proposed Work

The proposed STDM combines the two ideas of Steiner tree and the decentralization mechanisms and gives a better privacy protection when compared to some of the existing approaches. The Steiner tree structure of the network in WSN helps in efficient data transmission with minimum hops. It helps in achieving decreased communication overhead. The decentralization mechanism, applied to the data being transmitted, helps in privacy of data against privacy attacks where, an adversary cannot obtain the complete data and communication patterns even when it compromises one of the transmitting nodes. STDM approach avoids an adversary from tracing back to the source node. This paper gives the performance of STDM in data transmission with respect to path diversity, energy consumption, reliability, communication overhead and computation costs.

1.4 Organization of the Paper

Rest of the paper is organized as follows: Section 2 includes some of the related works; Section 3 details the proposed work; Section 4 illustrates performance analysis, and Section 5 deals with simulation and results; and Section 6 draws conclusions.

2 Related Work

A Steiner tree application in a datagram networks is given in [2] which aims at minimizing the number of nodes involved in routing over multicast tree and in maintenance. It proves that the cost of the created multicast tree is not necessarily higher than the cost of the trees created by other algorithms. One of the Steiner tree application in security is given in [4]. The security is achieved using keys for each sub-tree. When a key is given, a node can access the data only if its key matches with the subtree key given. It is proved that an adversary cannot obtain the key by intercepting the messages from the sensor nodes. It also briefs the Steiner tree construction which will be discussed in later section. There are many mechanisms available for privacy protection in WSN. A secure decentralized data transfer against node capture attacks for WSN is given in [12] to protect the privacy of the data against node capture attacks using secret sharing scheme. The experimental result shows that the node compromise ratio and the overhead are reduced in this approach. It was considered two types of emergent events: node fault and node capture. If a node fault occurs, the data cannot be transmitted to any of its neighboring nodes.On the other hand, if a node is captured by an adversary, it can be manipulated in any number of ways. It is assumed that the captured node fabricates the transmitted data. Privacy-preserving data aggregation in wireless sensor networks given in [7] states two mechanisms CPDA(Cluster-based Private Data Aggregation) and SMART (Slice-Mix-AggRegaTe) focusing on data privacy. CPDA and SMART use data-hiding techniques and encrypted communication to protect data privacy. Both CPDA and SMART result in efficiency with respect to privacy preservation, aggregation accuracy and computational overhead but the communication overhead is more in SMART. A scalable and distributed security protocol for multicast communication is proposed in [16] which explains how the sub-group keys and the node keys are organized and how the member join and leave are handled. It addresses the computational overhead in terms of key generation and encryption/decryption. It also analyses the communication overhead, message size, and storage overhead. When compared to existing protocol, the proposed protocol enhances the group performance especially in terms of computation and communication overhead at leave operations.

Secure and privacy-preserving data aggregation in WSN is proposed in [8] to protect the privacy of individual sensor readings. This approach results in better accuracy of data along with privacy preserved, each pair of sensor nodes shares secret key and encrypts messages when individual sensors report their privacy-sensitive data to protect the data privacy. In [17], achieving network level privacy in WSN for the data comes along with reliability and modest cost of energy as well as memory. It incorporates basic design features from related research fields such as geographic routing and cryptographic systems. Two notions used in this paper are: direction and trust. Both these notions (direction and trust) are used to provide reliable (non-malicious and non-faulty) secure paths for achieving robust route privacy. Direction information will help to forward packet to the destination in a timely manner and trust will help to forward the packets via reliable nodes. The experimental results of this approach gives an efficiency in terms of energy and memory usage. Node-failure Tolerance of Topology in WSNs is described in [25]. It explains the concept of node-failure along with the suitable topology which is efficient in tolerating node-failure. It also gives the mathematical analysis of the tolerance in terms of fault and intrusion and also the rules of fault and intrusion tolerance with head ratio of hierarchical topology are achieved. It also gives the mathematical analysis of the tolerance in terms of fault and intrusion and also the rules of fault and intrusion tolerance with head ratio of hierarchical topology are achieved.. It also gives the mathematical analysis of the tolerance in terms of fault and intrusion and also the rules of fault and intrusion tolerance with head ratio of hierarchical topology are achieved.

3 Proposed Scheme

STDM uses Steiner tree structure for the WSN and a decentralization mechanism for the data transmission. This section discusses the working of Steiner tree construction, decentralization and STDM along with their algorithms.

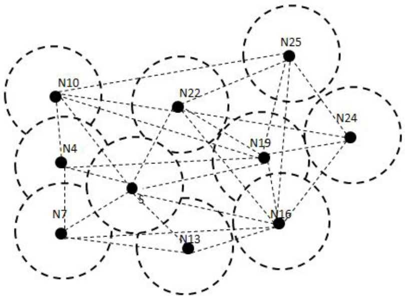

Figure 1: Network model

3.1 Steiner Tree Construction

Consider a WSN, V=(G,E), where G is a set of nodes and $E \subseteq V^2$ is the set of communication links, as shown in Figure 1. Let S be the source node. There exists a pair $(x, y) \in E$ i.e., x is able to communicate with y. Hence, the neighborhood set of x, $N(x)$, is said to contain all the nodes except x that have an edge or a communication link with x.

$$N(x) = y \neq x \wedge (x, y) \in E. \tag{4}$$

It is assumed that each node gets its location information accurately by some location services like GPS or any other positioning systems and there is a basic geographic routing protocol for the network. Each node has its own Node-state Information Table 1 containing fields like, PID (global identification of each node in the network), DID (location information of a node or Dynamic Identification), Cluster Head Flag (1 if it is a cluster head and 0 for a member node), Height Value (length of the path to the source node in a Steiner tree), Membership Flag (0 if a cluster has no cluster member and 1 otherwise) and Father Node (local routing information about the next node it selects for the transmission). Steiner tree is constructed using the location of each node considering the nearest nodes. A source node constructs a Steiner tree using nodes in its database, a node can participate in communication only if it is included in the database. An adversary cannot compromise a node without its PID and DID being present in the database.

Table 1: Node-state information table

PID	DID	ClusterHead Flag	Height Value	Membership Flag	Father Node
Node1	(11,20)	1	4	1	(1,9)

The five phases in Steiner tree construction are:

Algorithm 1 Algorithm for a Steiner tree construction

```
Begin
DID- location information of a node
S- source node
CH- Cluster Head
MN- Member Node
$R_{node_i d}$- radio range of a node
initialize node-state information table at source node S
repeat
    if $DID < R_S$ then
        Member node of S
        set $MN = 1$
    else
        if $DID > R_S,\ DID\ of\ node < DID_i$ then
            $CHF = 1$    //Cluster Head Flag
        else
            if $DID < R_{CH_i}$ then
                $MNF = 1$    //Member Node Flag
            else
                $CHF = 1$
            end if
        end if
    end if
until no more activities
end
```

1) Node-state information gathering phase: a node sends the join request to S along with its location information.The node S saves these information in its database and uses it to construct a Steiner tree. In order to prevent a malicious node from joining the network, a source node should verify each node. Authentication of each node is checked by a pre-shared key. Each node sends location information and $HMAC$ to the source node which includes the time T of sending the message.When a source receives the message, it validates T with the local current time *Clock*. I the inequality $|Clock - T| < \Delta t$ holds, then it accesses the the key from the database according to the *PID*. Then the source node calculates the value of $HMAC$ and compares it with the hash value in the received message. If they are equal then it proceeds to the next step, else the message is rejected.

2) Steiner tree construction phase: S constructs a Steiner tree depending upon the location information gathered. Steiner tree construction is represented in Algorithm 1.

3) Steiner sub-tree distribution phase: the obtained Steiner tree is divided into sub-trees for ease of data transmission [4]. The Steiner tree constructed along with sub-tree distribution is given in Figure 2, where, the nodes are represented along with their height values, and edges are represented with their sub-tree IDs. Example, $N22_{(1)}$ means the node $N22$ is one hop away from the source and belongs to the 3^{rd} sub-tree.

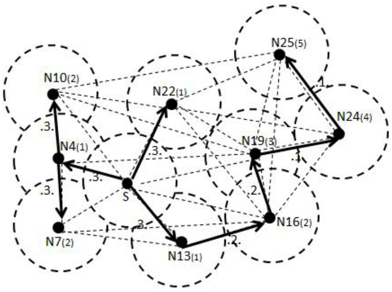

Figure 2: Steiner tree

4) Data delivery phase: the data is transmitted to the sink depending upon its location using unicast, multicast and broadcast technologies.

5) Steiner tree maintains phase: the tree is checked for node failures and new nodes to update the tree.

3.2 Decentralization Mechanism

The decentralization mechanism is given in Algorithm 2 where, source S partitions the data D and forwards these partitions through its neighboring nodes which route to the sink T. When these partitions reach T, it decrypts them after the number of partitions received is equal to some threshold value t sent by S. It then checks these decrypted data with the Cyclic Redundancy Check for errors.

Algorithm 2 Algorithm for a decentralization mechanism

```
Begin
S- source node
T- sink node
D- original data
t- threshold value given by S
$t_T$- number of partitions received at T
CRC- Cyclic Redundancy Code
D'- decrypted data at sink
CRC'- decrypted CRC
input: P //Number of neighboring nodes of source node (n-number of partitions)
output: D' //decrypted original data
get CRC value at S
at S:
apply Equation (2) for n partitions
at T:
if $t_T$==t then
    apply Equations (3) and (4) for $t$ partitions collected
    decrypt CRC value
    check D' with CRC' to get accuracy of D'
else
    ignore the data
end if
end
```

3.3 Steiner Tree-based Decentralization Mechanism (STDM)

STDM applies Steiner tree and decentralization mechanisms over WSN. Algorithm 3 shows the working of STDM. The STDM applied for Figure 1 is given in Figure 3. The source node S first collects the location information of all the nodes and constructs its Steiner tree. The node S sends a data D to the node *N25*, by partitioning it into *D1* and *D2*, through its neighboring nodes *N22* and *N13* respectively by encrypting them using Equation 1 along with the encrypted CRC code. These nodes use their Steiner structures to transmit the data. At node *N25* the received number of partitions is checked with the threshold value obtained from S and decrypted using Equation 2 and 3. The decrypted data is then checked with the decrypted CRC value to obtain the data accuracy. The STDM achieves the following important advantages:

Algorithm 3 Algorithm for a Steiner tree-based decentralization mechanism

```
Begin
S: source node
D: original data
P: number of neighboring nodes of S
$P_i$: $i^{th}$ neighboring node of S, $i = 1, 2, ...P$
$D_i$- partitions of D
T: sink node
call Steiner tree construction at S
call decentralization mechanism at S
repeat
    call Steiner tree construction for all $P_i$
    transmit $D_i$ to T through $P_i$
until no more events
call decentralization mechanism at T
end
```

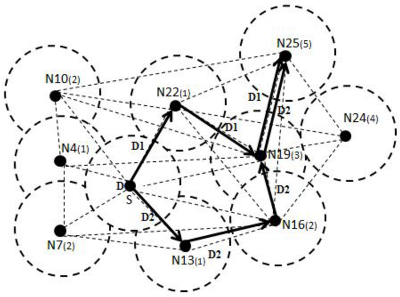

Figure 3: STDM for privacy protection

1) The path length of each node through which the data partitions are sent, varies among each other. Each node has its own path length to destination node which results in higher path diversities. An adversary cannot estimate the exact value of the path length needed to reach to the source.

2) Since there exist path variations and each neighboring node of a source node uses its own Steiner tree structure, it is difficult for an adversary to trace back the actual source node to obtain the original data. This is because, once a data is transmitted from the source node to its neighboring node, it uses the Steiner structure constructed from itself to the destination in order to transmit the data. So, even though the actual source node is S, the root of that particular path becomes the neighboring node and hence, when an adversary tries to trace back to the source node, it can reach only up to the neighboring node of the actual source node where it is not possible to get the original data.

3) If an adversary needs to compromise a node and join the network, it has to provide a location information and a private key which is common to all nodes. An adversary fails to provide the exact key and gets rejected from joining.

3.4 STDM and Privacy Attacks

1) *Monitoring and Eavesdropping*: Before gaining the access to any data from any node, the DID of the adversary should be present in the nodes under the

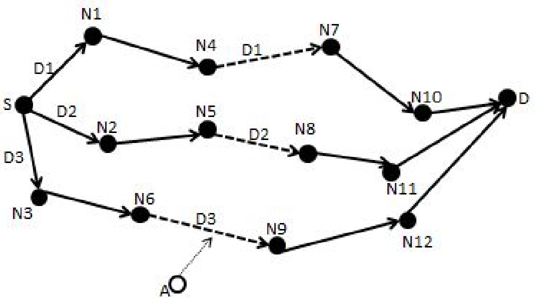

Figure 4: STDM and privacy attacks

Steiner structure. As given in Section 3.1, before construction of Steiner tree, nodes are checked in order to avoid malicious nodes from joining the network. Hence, when an adversary attacks through monitoring and eavesdropping, its GID cannot be approved which in turn reduces the possibilities of such an attack being successful. In case, an adversary obtains the data at a node, say D3 as shown in 4, it is an incomplete encrypted data. Since original data is not available at any node, the monitoring and eavesdropping attacks are failure to required extent.

2) *Traffic Analysis*: A source divides and distributes the data through its neighbor nodes. The neighbor nodes use the Steiner structure which starts from themselves to forward the data to the sink. Hence, a complete data can only be available at the source S. When an adversary tries to obtain the communication pattern, it should be one among the nodes that come under the transmission path i.e., its GID must be included in the group of nodes that come under the transmission path to the sink D through the respective neighbor node. In traffic analysis, the advantage of obtaining communication pattern is that the adversary can trace back to the source node of the respective transmission path to obtain the original data. But, it can only trace back to one of the neighbor nodes since the partition of the data are transmitted to D through the Steiner structure of a neighbor node which has an incomplete encrypted data. Hence, STDM reduces the success of traffic analysis.

3) *Camouflages Adversaries*: Even if an adversary compromise the node and obtains the data, it needs a key to decrypt it. The data is however a partition of the original data that can be obtained only if the number of partitions reached to D matches with the given threshold value. Hence, camouflages adversaries are made unsuccessful.

4 Performance Analysis

The performance of STDM is analyzed in terms of path diversity, energy consumption, reliability, communication overhead and computational cost.

4.1 Path Diversity

The path length of a node for each data partition differs with each other which results in path diversities. An adversary cannot estimate the value of the actual path length from any node. Hence, it is difficult for an adversary to trace back the actual source node to obtain the original data. The degree of the difference between paths that are actually taken by packets to the same destination injection into different nodes gives the path diversity of a network [13]. It depends on the data transmission mechanism used and utilization of the paths.

Let S be a source node and T be a destination node. Let N_i be the neighboring nodes of S to transmit the data D to T. Such a node N_i provides a path not containing an edge (N_i, S) and S itself. Let, the path length i.e., cost to a sink from N_i be C_i where $i = 1, 2, ..., n$ and n is the number of data partitions made. The optimal cost over all neighbors is formulated as $min(C_i)$. Let, $m \in N_i$ be the neighbor minimizing the length or cost C_i. To measure the path diversity, for each node $S \in V$, N_i nodes determine the set of costs C_i. In STDM, the source S uses all its neighbors to forward the decentralized data. Each N_i can use path towards T not containing the link connected to S. The degree of the difference between C_i's where $i = 1, 2..., n$ gives the path diversity of STDM and the value being higher guarantees the better privacy of the data transmitted.

4.2 Energy Consumption

In a Steiner tree, the same node may be used by different data partitions in order to reach the sink via shortest path and hence, increasing the energy consumed over those nodes. Each node plays different role in the network, like cluster head or member node, depending upon the path chosen for the transmission. The energy consumption increases with the number of destinations and their locations.

A sensor node utilizes its energy to carry out the three important functions as given below [20, 5]:

1) Acquisition: The energy consumed depends upon the monitoring carried out and is negligible.

2) Communication: Consumes more energy in terms of transmission and reception.

3) Data Processing: Very low energy is consumed for computations when compared to communication energy.

Each cluster head in STDM expends energy while transmitting as well as receiving a data from other cluster heads and also while receiving data from member nodes. Let E be the initial energy of each node and E_{ij} be the energy left at each node and N_{ij} be the node which transmits the data where, $j = 1, 2, ..C_i$. Let ET_{ij} be the energy consumed at a node for a data transmission of c bits over a distance m meters. Then ET_{ij} can be represented as

$$ET_{ij}(c, m) = c(E_{ec} + e_{amp} * m^{\alpha}), \tag{5}$$

where E_{ec} is the energy consumed during transmission, e_{amp} is the amplification and m^{α} is the transmitter/receiver distance. The energy consumed to receive a data, $ER_{ij}(c)$, is given by,

$$ER_{ij}(c) = c * E_{ec}. \tag{6}$$

The total energy consumed,E_{total} is the sum of energy consumed by the sub-trees,E_{st}, along the path to the sink [15]. It is given by,

$$E_{total} = \sum_{st=1}^{j} E_{st}, \tag{7}$$

where, E_{st} amounts to the total energy consumption of the data flows whose number is the number of nodes, n'. The number of hops in each flow of interest, l_i, depends on the nodes' locations. It is given by

$$E_{st} = \sum_{i=1}^{n'_i} \sum_{j=1}^{l_i} E_{trs}, \tag{8}$$

where, E_{trs} is the energy consumed by a transmitter, receiver and a sensor which in turn gives the energy consumed for a single hop. Hence, E_{trs} is given by,

$$E_{trs} = ET_{ij} + ER_{ij} + E_s. \tag{9}$$

The energy consumption of STDM depends upon the role of a node in the Steiner tree i.e., a Cluster Head of Steiner structure of one neighbor node may be a member node in the Steiner structure of other neighbor node. It may or may not play the same role for multiple nodes. Energy consumption varies with the role of a node in the tree i.e., it increases for a Cluster Head and decreases for a member node.

4.3 Reliability

In STDM, if an adversary needs to compromise a node and join the network, it has to provide DID and a private key which is common for all nodes. An adversary fails to provide the exact key and gets rejected from joining. The probability of the data D being compromised, i.e., P, can be obtained by the ratio of number of nodes that collect all dispersed shares of each node to the sink node, N_p, to the number of combinations of nodes compromised, N_c. Lesser the value of P more the scheme is reliable, i.e., the reliability R of STDM is inversely proportional to P. That is,

$$R \, \alpha \, \frac{1}{P}, \tag{10}$$

where,

$$P = \frac{N_p}{N_c}. \tag{11}$$

The number of combinations of nodes,N_c, can be computed using

$$N_c = \sum_{l=0}^{N} nC_l = \sum_{l=0}^{N} \frac{N!}{l!(N-l)!} \tag{12}$$

where, N is the total number of sensor nodes and l is the number of nodes compromised. By calculating the value of P using Equations (11) and (12), the reliability of the proposed scheme is obtained. Due to the Steiner structure and decentralization mechanism used, an adversary cannot obtain the exact data or communication pattern of STDM. Hence, even if the node compromise ratio increases, the scheme is reliable in terms of privacy of D.

4.4 Communication Overhead

Number of bits of data added to the packet to carry routing information, operational instructions and check code consuming the available resource of a node gives the communication overhead [1]. Consider, a Minimum Spanning Tree(MST) and a Steiner tree, as shown in Figure 5. Since the number of edges between any two nodes is less in Steiner tree compared that in MST, its communication overhead decreases when compared to that of MST structured networks. The decentralization mechanism also reduces the communication overhead of each node participating in the data transmission. In STDM, the communication overhead per node for the transmission of a particular data partition, C_{ij} where $i = 1, 2, ..., n$ and $j = 1, 2, ...C_i$, of a node is proportional to the number of packets that are being forwarded and the bandwidth [22]. i.e.,

$$C_{ij} = \frac{n_{ij}}{BW_{ij}}, \tag{13}$$

where, n_{ij} is the number of packets forwarded by a node N_{ij} and BW_{ij} is the bandwidth available.

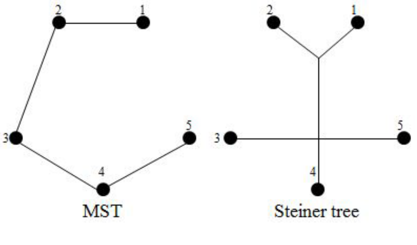

Figure 5: Difference between steiner tree and minimum spanning tree

4.5 Computational Cost

Computation cost depends on the amount of computations done at a node. In STDM, computations are done at each node participating in a data transmission.

1) At source node,

   a. The data is divided into partitions depending upon the number of neighboring nodes available.

   b. Each data partition is encrypted and transmitted along with encrypted check code.

2) Each node in the transmission, including source node, keeps track of energy consumed.

3) At sink node,

   a. The number of partitions arrived must be counted and be checked whether it is equal to t.

   b. Once t partitions are available, it starts decrypting them to get the original data.

   c. Compare the decrypted data with the decrypted check code to get the accuracy.

STDM uses Lagrange's Interpolation method for encryption and decryption. The source node S and the sink node T encrypts and decrypted the data partitions respectively. According to Equation (1), S has to do $O(t+1)$ operations for encryption i.e., $O(t)$. The decryption (Equation (2)) of data at T needs $O(t + 1\lambda_t)$ operations where λ_t is a constant and hence, it needs $O(t)$ operations. All the nodes including S and T keeps track of energy consumed. It needs $O(n)$ operations [11]. The transmission and reception cost is $O(n)$. So, nodes S and T do $O(t + n + n) = O(t + n)$ operations. The intermediate nodes do $O(n + n) = O(n)$ operations. Since, only S and T do encryption and decryption respectively and intermediate nodes participate only in transmission and reception due to the decentralization mechanism which guarantees the privacy of the data, the communication cost of STDM is less.

5 Simulation and Results

For simulation purpose we assume the communication range of each node as 50m, the area of simulation of 500*500m, the density as 10 nodes per communication range, and a total of 300 nodes are deployed. Each node has equal transmission power level. Simulation time is 1800 seconds. Node placement strategy is random. Bandwidth is considered to be 200Kbps.

Higher the path diversity, better the privacy provided in WSN. Consider a Steiner tree with 6 paths to the sink. Let the data D be partitioned into 6 parts $(D1, D2, D3, D4, D5, D6)$. The transmission of these partitions through the Steiner tree using 6 neighboring nodes of the source S $(N1, N2, N3, N4, N5, N6)$ gives different path lengths which results in high path diversities when compared to the other two existing approaches PSR-hop (Phantom Single-path Routing with hop-based approach) and PSR-sec (Phantom Single-path Routing with sector-based approach) [10] as shown in Figure 6. This makes it difficult for an adversary to estimate the exact path length and to trace the source node and hence, provides the privacy of the sensed data.

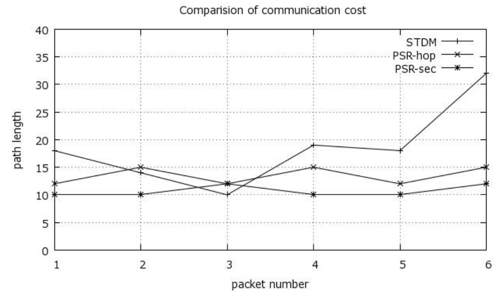

Figure 6: Path diversity in privacy schemes

The energy consumption for the decentralization of data over Steiner tree can be analyzed as follows:

1) Because of the Steiner structure of the network, most of the data partitions may use the same nodes for transmission for obtaining shortest path. This will slightly increase the energy consumed over those nodes of the network.

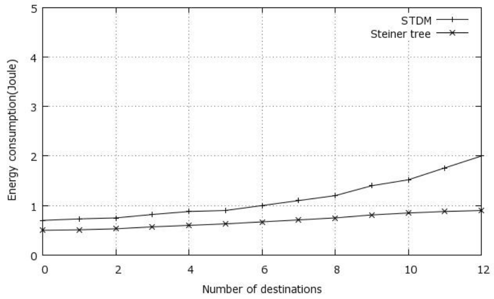

Figure 7: Energy consumption of proposed scheme

2) The cluster head node in Steiner tree of one neighboring node may be a member node in that of other neighboring node. This will make the energy consumption being shared by different nodes at each path.

3) The energy consumption increases with the number of destinations and their locations.

This analysis for STDM is compared with one of the existing approaches of Steiner tree shown in Figure 7.

The reliability of STDM, is compared with that of the decentralization mechanism applied without Steiner tree structure (DWST) and noticed that the possibilities of node compromising decreases in the proposed approach as shown in Figure 8.

This is because, in a Steiner tree any node can transmit data to a node only if its DID is present in the database which means that the node is a part of the constructed Steiner tree. In STDM, even if an adversary obtains the data it is incomplete because of decentralization. Hence, the proposed approach is considerably reliable and reduces data compromise.

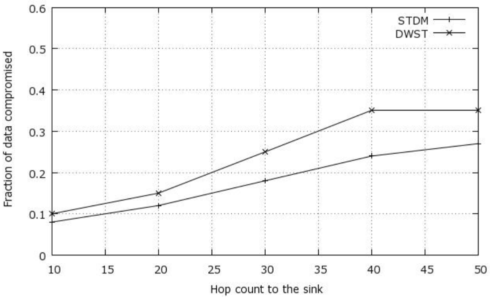

Figure 8: Reliability of STDM

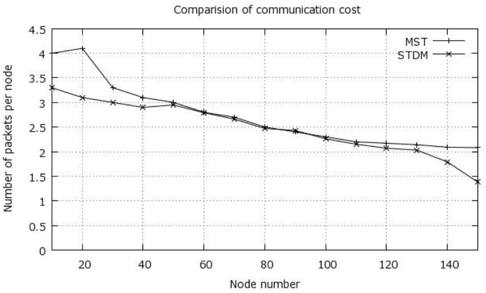

Figure 9: Communication overhead of STDM

The communication overhead of STDM is compared with that of a network with a MST structure as shown in Figure 9. The Steiner tree structure is said to have minimum number of edges between any to nodes compared to MST which reduces the communication overhead over WSN. The decentralization mechanism helps in reducing the communication overhead of nodes. Hence, STDM has a reduced communication overhead compared to a network with MST structure.

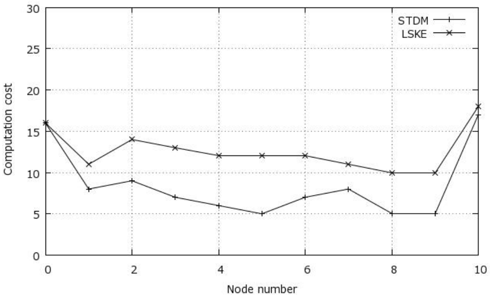

Figure 10: Computation cost of STDM

The computation cost of STDM is measured for getting the count of the number of partitions of data to be transmitted to be made, encryption and decryption of these partitions. This cost is compared with that of a key establishment scheme [3] for WSN, which includes encryption and decryption processes at all the nodes. Considering the computations for one data partition transmission in STDM, Figure 10 shows that the proposed scheme is more efficient in terms of computation cost compared to that of Key establishment scheme. It is because STDM encrypts the data at S and decrypts the data at T, the data is safe at intermediate nodes because of decentralizing as well as Steiner tree structure.

6 Conclusion

This paper presents the performance analysis of a Steiner Tree-based Decentralization Mechanism (STDM) for Privacy Protection in WSN, which mainly aims at the privacy of sensed data and data transmission along shortest path. STDM achieves a better performance over protecting privacy of the sensed data when compared to some of the existing mechanisms. The Steiner tree structure of the network results in minimum number of nodes for data transmission and the decentralization mechanism protects the data against privacy attacks. The decentralization mechanism over Steiner tree results in high path diversities and hence, reduces the possibilities of occurrence and success of the privacy attacks. The simulation results show that STDM is efficient in terms of energy consumption, reliability, communication overhead and computation cost when compared to the existing security approaches for WSN.

References

[1] "Overhead,". Business Dictionary.

[2] E. Aharoni and R. Cohen, "Restricted dynamic steiner trees for scalable multicast in datagram networks," *IEEE/ACM Transactions on Networking*, vol. 6, pp. 286–297, June 1998.

[3] A. K. Das, "A location-adaptive key establishment scheme for large-scale distributed wireless sensor networks," *Journal of Computer*, vol. 4, no. 9, pp. 896–904, 2009.

[4] R. Fan, J. Chen, J. Q. Fu, and L. D. Ping, "A steiner-based secure multicast routing protocol for wireless sensor network," in *Second International Conference on Future Networks*, pp. 159–163, 2010.

[5] A. O. Fapojuwo, C. K. Tse, and F. C. M. Lau. "Energy consumption in wireless sensor networks under varying sensor node traffic,", 2010.

[6] M. Gruteser, G. Schelle, A. Jain, R. Han, and D. Grunwald, "Privacy-aware location sensor networks," *USENIX Workshop on Hot Topics in Operating Systems (HotOS IX)*, p. 28, 2003.

[7] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and F. T. Abdelzaher, "Pda: Privacy-preserving data aggregation in wireless sensor networks," in *26th IEEE International Conference on Computer Communications*, pp. 2045–2053, 2007.

[8] W. He, H. Nguyen, X. Liu, K. Nahrstedt, and F. T. Abdelzaher. "Spda: A secure and privacy-preserving data aggregation in wireless sensor networks,". Tech. Rep. 11275, Nov. 2006.

[9] F. K. Hwang, D. S. Richards, and P. Winter, *Steiner Tree Problem.* ELSEVIER, 1992.

[10] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," in *Proceedings of 25th IEEE International Conference on Distributed Computing Systems*, pp. 599–608, Columbus, OH, USA, 2005.

[11] N. Kogan, Y. Shavitt, and A. Wool, "A practical revocation scheme for braodcast encryption using smartcards," *ACM transactions on Information and Security*, vol. 9, no. 3, 2006.

[12] E. Kohno, T. Ohta, and Y. Kakuda, "Secure decentralized data transfer against node capture attacks for wireless sensor networks," *Autonomous Decentralized Systems (ISADS'09)*, pp. 1–6, 2009.

[13] Pascal Merindol and Antoine Gallaise, "Path diversity in energy-efficient wireless sensor networks," *20th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 2280–2284, 2009.

[14] G. Padmavathi and D. Shanmugapriya, "A survey of attacks, security mechanisms and challenges in wireless sensor networks," *International Journal of Computer Science and Information Security*, 2009.

[15] W. Y. Poe and J. B. Schmitt, "Node deployment in large wireless sensor networks: Coverage, energy consumption, and worst-case delay," *Asian Internet Engineering Conference (AINTEC)*, pp. 77–84, 2009.

[16] I. A. Saroit, S. F. El-Zoghdy, and M. Matar, "A scalable and distributed security protocol for multicast communications," *International Journal of Network Security*, vol. 12, pp. 61–74, Mar. 2011.

[17] R. A. Shaikh, H. J., Brian J. D.AUriol, H. Lee, and S. L. Y. J. Song, "Achieving network level privacy in wireless sensor networks," *Sensors(2010)*, vol. 10, no. 3, pp. 1447–1472, 2010.

[18] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[19] S. Sharma. "Energy-efficient secure routing in wireless sensor networks,". Tech. Rep. 1515, June 2009.

[20] F. Shebli, I. Dayoub, A. O. M'foubat, A. Rivenq, and J. M. Rouvaen. "Minimizing energy consumption within wireless sensors networks using optimal transmission range between nodes,", 2007.

[21] Scribe Siddharth, Barman Lecturer, Shuchi Chawla, and Steiner Tree. "Topic: Steiner tree; greedy approximation algorithms date: 01/25/07,".

[22] J. Sonnek, J. Greensky, R. Reutiman, and A. Chandra. "Starling: Minimizing communication overhead in virtualized computing platforms using decentralized affinity-aware migration,", 2010.

[23] M. Tehranipoor. "Steiner tree problem,". tech. rep., ECE Department, 2008.

[24] J. P. Walters, Z. Liang, W. Shi, and V. Choudhary, *Wireless Sensor Network Security: A Survey.* Boca Raton: Auerbach publications:Security in Distributed, Grid, and Pervasive Computing, 2006.

[25] L. M. Wang, J. F. Ma, and Y. B. Guo, "Node-failure tolerance of topology in wireless sensor networks," *International Journal of Network Security*, vol. 7, pp. 261–264, Sep. 2008.

B. Sathish Babu received his bachelors and masters degree in computer science and engineering from Bangalore university, completed his Ph.D. in ECE at Indian Institute of Science, Bangalore, during 2009. His research interests Security in mobile communication; Cognitive agents based control solutions for Networks, Grid Computing, Cloud Computing Scheduling and Security Issues, Context-aware Trust issues in Ubiquitous Computing, Privacy issues in WSN, and so on. He wrote 15 international conference papers and 10 international journal papers in his area of research. Published a book on Mobile and Wireless Network Security(TMH) in the year 2010. Dr. B. Sathish babu, has given more than 25 invited talks in national level workshops and FDP on various topics of current trends.

Jayashree N received her bachelors and masters degree in computer science and engineering from Visvesvaraya Technological university. Her research interests Security in data transmissions, data privacy in wireless sensor networks and key management techniques. She is the beginner in this field and published 1 international paper on one of the novel technique to solve privacy issues in 2011.

Pallapa Venkataram received his Ph.D degree in Information Sciences from the University of Sheffield, U.K.in 1986. He is currently a Professor of Electrical Communication Engineering with Indian Institute of Science, Bangalore, India. His research interests includes protocol engineering, wireless networks, network management, computational intelligence applications in communication, mobile computing security and multimedia systems. He is a Fellow of IEE (England), Fellow of IETE(India) and a Senior member of IEEE Computer Society. Dr. Pallapa is the holder of a distinguished visitor diploma from the Orrego University, Trujillo, Peru. He has published over 150 papers in International/national Journals/conferences.